

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2003, 13:15:55 ; Search time 74 Seconds

(without alignments)
1623.730 Million cell updates/sec

Title: US-09-494-297-2

Perfect score: 3945

Sequence: 1 MKKTRFPKMLTMTQRYLS.....TAGISLGIWGIHTIRKRD 757

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq.19Jun03.*

```
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3945	100.0	762	23 ABP29047	Streptococcus poly
2	236.5	6.0	23	ABP29050	Streptococcus poly
3	181	4.6	20	AAW89421	Moraxella catarrha
4	180	4.6	597	21 AAY90257	Streptococcus equi
5	178	4.5	898	20 AAW89413	Moraxella catarrha
6	161	4.1	1416	20 AAY00211	Enterococcus faeca
7	161	4.1	23	ABP43430	E faecalis EF104 a
8	161	4.1	1416	24 ABU13709	Enterococcus faeca
9	161	4.1	1448	20 AAY00210	Enterococcus faeca

10	161	4.1	1448	23 ABP43429	E faecalis EF104 p
11	161	4.1	1448	24 ABU13708	Enterococcus faeca
12	159	4.0	1612	23 ABW47333	Listeria monocytog
13	156	4.0	1530	23 ABW47329	Listeria monocytog
14	155	3.9	832	23 ABW47336	Listeria monocytog
15	154	3.9	1849	18 AAW17900	Photobacterium lumini
16	152.5	3.9	2013	23 ABW47751	Listeria monocytog
17	152	3.9	1849	19 AAW55573	Toxin TcdA11, enco
18	152	3.9	2516	18 AAW1899	Photobacterium lumini
19	152	3.9	2516	19 AAW55572	Toxin TcdA, encode
20	152	3.9	2516	22 AAB72609	Photobacterium tcdA
21	152	3.9	2516	23 ABG32651	P. luminescens (W-
22	152	3.9	2517	22 AAB72611	Modified Phototrab
23	152	3.9	2537	22 AAB72614	TcdA toxin-zeln ER
24	150	3.8	1161	23 AAE22273	Streptococcal fibr
25	149	3.8	1315	23 AAW08642	S. aureus SdrD pro
26	149	3.8	1315	24 ABU18969	Pathogen specific
27	149	3.8	2032	20 AAY00238	Enterococcus faeca
28	149	3.8	2032	20 AAY00240	Enterococcus faeca
29	149	3.8	2032	20 AAY00242	Enterococcus faeca
30	149	3.8	2032	23 ABP43457	E faecalis EF123 p
31	149	3.8	2032	23 ABP43459	E faecalis EF124 p
32	149	3.8	2032	23 ABP43461	E faecalis EF125 p
33	149	3.8	2032	24 ABU13736	Enterococcus faeca
34	149	3.8	2032	24 ABU13738	Enterococcus faeca
35	149	3.8	2032	24 ABU13740	Enterococcus faeca
36	147	3.7	1349	22 AAW34402	Staphylococcus aur
37	147	3.7	1349	22 AAW34404	Staphylococcus aur
38	147	3.7	1563	22 ABB58432	Staphylococcus aur
39	146	3.7	1185	13 AAR22675	Drosophila melanog
40	145	3.7	1112	20 AAY08603	Collagen binding p
41	144.5	3.7	940	23 ABB47334	S. pyogenes SFRP-
42	144.5	3.7	1092	19 AAW41602	Listeria monocytog
43	144.5	3.7	2261	24 ABU18914	Staphylococcus epi
44	144.5	3.7	2283	24 ABB56876	Pathogen specific
45	144.5	3.7	2244	22 AAW37120	Staphylococcus epi

ALIGNMENTS

RESULT 1	
ABP29047	standard. Protein; 762 AA.
ID	ABP29047
AC	ABP29047;
DT	02-JUL-2002 (first entry)
DE	Streptococcus polypeptide SEQ ID NO 7270.
XX	
KW	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW	group A streptococcus; Streptococcus pyogenes; antibacterial;
KW	antiinflammatory; Infection; vaccine; meningitis; gene therapy.
OS	Streptococcus pyogenes.
XX	
PN	WO200234771-A2.
PD	02-MAY-2002.
PF	29-OCT-2001; 2001WO-GB04789.
PR	27-OCT-2000; 2000GB-0026333.
PR	24-NOV-2000; 2000GB-0028727.
PR	07-MAR-2001; 2001GB-0005640.
PA	(CHIR-) CHIRON SPA.
PA	(GENO-) INST GENOMIC RES.
PI	Telford J, Maignani V, Margalit Ros YI, Grandi G, Fraser C;
PI	Tetelin H;
XX	

DR WPI: 2002-352536/38.
 DR N-PSDB: ABN69678.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1; Page 3879; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 XX
 SQ Sequence 762 AA;
 Query Match 100.0%; Score 3945; DB 23; Length 762;
 Best Local Similarity 100.0%; Pred. No. 2,4e-285;
 Matches 757; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKKTREPKNLTINTQVLSKNSKRTVTLVGVFLMIFALVTSVAGAKYEGLVESSTPN 60
 DB 6 MKKTREPKNLTINTQVLSKNSKRTVTLVGVFLMIFALVTSVAGAKYEGLVESSTPN 65
 QY 61 AINPDSSEYRWYGYESYVGRHPYKQFVAHDLRVNLEGSRSYQVYCFNLKRAFLGSD 120
 DB 66 AINPDSSEYRWYGYESYVGRHPYKQFVAHDLRVNLEGSRSYQVYCFNLKRAFLGSD 125
 QY 121 SSYKRWYKKHDKISTKEDYAMSRLTGDENQKLRVAVNGHQNANGIMEGLEPINA 180
 DB 126 SSYKRWYKKHDKISTKEDYAMSRLTGDENQKLRVAVNGHQNANGIMEGLEPINA 185
 QY 181 RVTQEAWYVSDNAPISNPDESFRSESNLVSTSLKROALKOLIDNLTAKMKQY 240
 DB 186 RVTQEAWYVSDNAPISNPDESFRSESNLVSTSLKROALKOLIDNLTAKMKQY 245
 QY 241 PDDFQSLFSESEDKGDKYNGYONLSGLVPTKPPPGDPMPMPNOPTTSVLRKYAI 300
 DB 246 PDDFQSLFSESEDKGDKYNGYONLSGLVPTKPPPGDPMPMPNOPTTSVLRKYAI 305
 QY 301 GDYSKLEGGATLQLTGNVNSFOARVSSNDIGERITLSGTYTLTFLNSPAGYSTAEP 360
 DB 306 GDYSKLEGGATLQLTGNVNSFOARVSSNDIGERITLSGTYTLTFLNSPAGYSTAEP 365
 QY 361 TFFVEAGKYVTIIDKQIENPNKEIYVPEYVEAVNDEEFSVLTTOYAKFYAKNNGS 420
 DB 366 TFFVEAGKYVTIIDKQIENPNKEIYVPEYVEAVNDEEFSVLTTOYAKFYAKNNGS 425
 QY 421 SQVYVCFNADLKSPDSEDEGKTMTDPTTGEVKKYTHIAGRDLEFKYTVKPRDTPDTFLK 480
 DB 426 SQVYVCFNADLKSPDSEDEGKTMTDPTTGEVKKYTHIAGRDLEFKYTVKPRDTPDTFLK 485
 QY 481 HIKVIEKGYREGGALEYSGLTETQLRATQALAIYFTDSAELEDKLKDYGFGDMND 540
 DB 486 HIKVIEKGYREGGALEYSGLTETQLRATQALAIYFTDSAELEDKLKDYGFGDMND 545
 QY 541 STLAVAKIIVEYADSNPQTLDPFIPNNNKYQSLIGQWHPEDLVDIRREDKKEVI 600
 DB 546 STLAVAKIIVEYADSNPQTLDPFIPNNNKYQSLIGQWHPEDLVDIRREDKKEVI 605

QY 601 PVTHNLTLRKTYTGLADRRKDFEHEITLKNKQELLISQVTKDKTNLEKDGKATINLK 660
 DB 606 PVTHNLTLRKTYTGLADRRKDFEHEITLKNKQELLISQVTKDKTNLEKDGKATINLK 665
 QY 661 HGESLTQGLPEGSYLVKKEFDSRGYKVKNSQVANAATYSKGTSDFLAPENKEPV 720
 DB 666 HGESLTQGLPEGSYLVKKEFDSRGYKVKNSQVANAATYSKGTSDFLAPENKEPV 725
 QY 721 VPTGVDPKINGYALAIYIAGISLGIMGIIHTRIRKHD 757
 DB 726 VPTGVDPKINGYALAIYIAGISLGIMGIIHTRIRKHD 762
 RESULT 2
 ABP29050
 ID ABP29050 standard; Protein; 340 AA.
 XX
 AC ABP29050;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 7276.
 XX
 KW Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus pyogenes.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Maignani V, Margait Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 DR WPI: 2002-352536/38.
 DR N-PSDB: ABN69681.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1; Page 3880; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 XX
 SQ Sequence 340 AA;

ID	AAW89413	standard; Protein; 898 AA.
AC	AAW89413:	
AD	21-JUN-1999	(first entry)
DE	Moraxella catarrhalis	lactoferrin binding protein 2 (Lbp2).
KM	lactoferrin receptor; lactoferrin binding protein; Lbp2;	
RN	LbpB gene; infection; otitis media; sinusitis; conjunctivitis;	
RV	pneumonia; bronchitis; tracheitis; emphysema; diagnosis; therapy;	
VZ	vaccine; Branhamella catarrhalis.	
XX	Moraxella catarrhalis.	
XX	Key	Location/Qualifiers
FT	Misc-difference	632
FT	/note= "encoded by AAR"	
FT	Region	430..435
FT	/note= "conserved epitope"	
PX	MO9855606-A2.	
PD	10-DEC-1998.	
PF	02-JUN-1998;	98WO-CA00544.
PR	08-MAY-1998;	98US-0074658.
PR	03-JUN-1997;	97US-0867941.
PA	(CONN-) CONNAUGHT LAB LTD.	
PI	Du R, Klein MH, Loosmore SM, Wang Q, Yang Y;	
DR	WPJ: 1999-070266/06.	
DR	N-PSDB; AAV82019.	
PT	Lactoferrin receptor genes from Moraxella,	especially M. catarrhalis
PT	- useful to diagnose Moraxella infection e.g. to detect	otitis media
PT	due to M. catarrhalis infection and to immunise against such	infections
PS	Claim 8; Fig 2; 202pp; English.	
CC	This protein comprises lactoferrin binding protein 2 (Lbp2) of	
CC	of Moraxella catarrhalis (Branhamella catarrhalis) 4223. It is	
CC	encoded by the lbpB gene of the lactoferrin receptor (lfr) locus	
CC	(see AAV72019) identified in the M. catarrhalis 4223 genome.	
CC	Immunogenic compositions, including vaccines, based upon expressed	
CC	recombinant lbpI and/or lbp2 and/or ORF3 proteins (see AAW89413-21),	
CC	portions of these, or their analogues, can be prepared for	
CC	prevention of diseases caused by Moraxella. M. catarrhalis is a	
CC	causative agent of otitis media and has been associated with	
CC	sinusitis, conjunctivitis and inflammatory diseases of the lower	
CC	respiratory tract, such as pneumonia, chronic bronchitis,	
CC	trachetitis and emphysema.	
SQ	Sequence	898 AA;
OY	Query Match	4.5%; Score 178; DB 20; Length 898;
Db	Best Local Similarity	18.6%; Pred. No. 0.00043;
Matches	171; Conservative	118; Mismatches 305; Indels 324; Gaps 42
OY	4 TRPNKLTNTLRVLSKNSRFFTVLVGVLFMTALVTSMVGAKTVEGLYESPPMAIN	63
Db	79 TTDPNGDNQLTFQ-----AOKTPAAGGFVM-----GKIDRTSPKN-D	115
OY	64 PDSSSE--YYRWYG--YESYVAGHP-----YYQFRVAHDLRVNL	98
Db	116 PDIYSNDLVQDQGKLIVGDIAHRPDCIGTGKRLROPIITANDIKPLPYFNKFALSDLDHS	175

QY	99	EGSRASYOVYOFNLKKAFLPLGSDSSVYK--WYKKHGISTGKPEDYAMSP-----RIT	147
Db	176	ERHR-EDPKKLNTIKVYGYGNLTTPSKNNTYINHQADNKKNNKPVDPYENIRFGYLELQ	234
QY	148	GDELINOK-----LRAVYNGHPONAGIMEGLEPLN-----	178
Db	235	GSLSLTQKNMADTPDXDKRIKRPMLILFYHG--ENASQSLPSAGKFTYGNMLYLSDVKKRP	292
QY	179	AIRVYDPAWVYSDNAPISNPDESEFKRESNLVSTQSLSL-----	219
Db	293	ALASPDREVGVYLTNAGSKSN-----EGDVSAAHILYINGEQYKHTPATYQVDEDTN	343
QY	220	-MGQALKQLIDPLMLAKMKQYVDDP-----QLSTFESEDK--GDKYKNGYGNLNLSSG	269
Db	344	SLTGKLSYSDNPNOQAQGYTKRSQDTTKKVENEDVQIDAKINGNFVGTALNVEN	403
QY	270	L-----VPTKPTPGDPMPPMPNOPTSVILIRKAYIAGYSKLEGGATQLGLGVNVS	321
Db	404	TETAPFIKELFKSKANPNP--NPN-----SDTLEGGFYGESEDEL--	442
QY	322	FOARVSSND-----IGERIELSDGTYLTLELNSPAGYIAEPITTFYVGAQKYTTIIDG	375
Db	443	--AGKELSDNMAASYVVGKRDCTOKPVATKIVYSAFGE--KSTSDVNETIGRIINS	498
QY	376	KOLENPKELIVE---PYSEAYENDP-----EESVYLTQN-----YAKFYAK	415
Db	499	KKLMDAVNEKIDMGDILPTSDERIDEPNGEKKAEFTTKYSSSTQAVPAVFGQDHKFTF--	556
QY	416	NKNGSSOVVYCFNADLKSPDSEDDGKMTPTDFTTGEV-----	453
Db	557	--NGNYDILSASSVDKLADAVKAKOSIKEKYPNATLNKNQVTAIVLEAKDNKPYTA	614
QY	454	-----KYNHIA-GRDLP-----KYTKPRDTPDPPFLKHKKIVY-----	488
Db	615	IRAKSYOHISFGETLYLNDANQPTRSYFVOGGRADTSTLLPRAKFTYNGLMAGYLIQKK	674
QY	487	EKGREKGOAIEYSG-----LTETOLRAATOLAITYFT-----DSAELEDKDLKDYHG	535
Db	675	DKGSSNNEELIKKKGHODYLTTFD-----FPREDDDDLTASDSSQDDDAHGD	722
QY	536	GDMNDSTLAVAKILVEYADDSNPQOLTDLDFPIPNNNKYQSLIGTQWHPEDLVDIRMED	595
Db	723	DDL-----IASDSDQDDADGDDSDDL--GDGADDAAAGKYVHAGN---IRPEF	767
QY	596	KKEVITYP-----THNLTL-----KTYVTGLAGRTKDFHFEILKNNKQELL	637
Db	768	ENKYLPLNEPTHEKRTALDGKMKAKEDVDFTNSLTGKLNERGVIYDI--KNGKIDGT	823
QY	638	SOYVKTDTKTULEFKDCKATINLKHGESLTLQGLPREGSYSLVKEFTSEGYKYKVSQVAA	696
Db	826	GFLAKADVPYRYREVG-----NNQGG-----GFLINIKIDIVKQGFPGTNGEBLAG	871
QY	697	-----NATVSKTG	704
Db	872	QLQYDKDGDGINDTAEKAG	889
RESULT 6			
AAV00211			
ID AAV00211 standard; Protein; 1416 AA.			
AAV00211:			
XX	20-APR-1999	(first entry)	
XX	Enterococcus faecalis antigenic polypeptide fragment EF104.		
XX	Enterococcus faecalis; infection; vaccine; immune response; diagnosis;		
XX	detection; attenuation; antigenic.		
XX	Enterococcus faecalis.		

```

PN  WO9850554-A2.
XX
XX  12-NOV-1998.
XX
XX  04-MAY-1998; 98WO-US08959.
XX
XX  14-NOV-1997; 97US-0066009.
XX  06-MAY-1997; 97US-0044031.
XX  16-MAY-1997; 97US-0046655.
XX
PA  (HUMA-) HUMAN GENOME SCI INC.
XX
XX  Bailey C, Choi GH, Hromockyj A, Kunsch CA;
XX  WPI; 1999-070095/06.
XX  N-PSDB; AAX20201.
XX
PT  New isolated Enterococcus faecalis polynucleotides - used to develop
PT  products for the detection of Enterococcus and for use in vaccines
PT  for prevention or attenuation of Enterococcus infection
XX
PS  Claim 9; Page 203; 301pp; English.
XX
XX  The present sequence represents an antigenic polypeptide fragment
XX  isolated from Enterococcus faecalis. The present invention describes
XX  genes, proteins and antigenic polypeptides isolated from E. faecalis.
XX  The proteins can be used in vaccines for preventing or attenuating an
XX  infection caused by a member of the Enterococcus genus in an animal.
XX  They can also be used for detecting Enterococcus antibodies in a sample.
XX  The nucleotide sequences can be used for detecting Enterococcus nucleic
XX  acids. Products from the present invention can also be used for
XX  screening compounds to identify agonists and antagonists of E. faecalis
XX  protein activity.
XX
SO  Sequence 1416 AA;

Query Match 4.1%; Score 161; DB 20; Length 1416;
Best Local Similarity 21.1%; Pred. No. 0.017;
Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;

QY  39 AUYTSMVGAKTIVGLVES---STPNAINPDSSEYMYGYESVVRGHPPYKKORVAHDLR 95
DB  484 SUSTPYIGPKAIQVSDOTIEPIISVNP-LNLETAMGNDO-----NGAYSSR 531
QY  96 --VNEGSRSYOYVCNFKKAP---LGSDDSVKKYKKHDSISTKFEDYAMSPRITCD 149
DB  532 TTYSVWGSKKEPIQNLKIKHNYLSLRATKEIYYK---LGT---DYTTPISDGS 584
QY  150 ELNOKLRVWYNGHPQNNANGIMEGL---PLNAIRVTOEAVWYSDNAPISNPDESFK 204
DB  585 VIFETTPITNEIQIPGFNVVPDPLPKDSIPVDTPITWSAGELPVPDVTVT---NSK 641
QY  205 RESESNLVSTQSLSLRQALQKLDINLTKMKQYPPDFQSLFSESDGDKYNNGYON 264
DB  642 RGSERTLQSSKNOFLVARNDSFSLSVRKIPAGA--DVLFDIYVSN--DOVDSIYPO 697
QY  265 LLSGGLVPTKPPTPGDPMP--PNOPQTSVLI-----RKYAIGYSKL---LEGATQ 313
DB  698 YMRGQYFPKPMPPNSPGYPTTFEDNTNSYTFDFPKTKRITII-EKKNANGIDVPTLY 756
QY  314 LIG-----DNVSFOARVSSNDIGERIELSDGYTTLFELNSPAGYSIAEPITFVEAG 367
DB  757 ITGTAKPEQSNNEGASASVQN---EALDIISAT-----QAANPLTKNVTKT 801
QY  368 KYVT-IIDCK-QIENPKIEIVPEVEANDFEESVLTQNYAKFYAKKNGSSQYV 424
DB  802 TVTTKNIDNKTHVKNKPTIELTPKGTNMOIDNSITV-----KGVPEDA 846
QY  425 YCFNADLKSPDSEDCGKTWPTFTGE---VKYTHIAGRLDKFYTVKPRDTPDFLKH 481
DB  847 Y-----SLEKTKNKAIVFKRYTILENTIIEYNVSANAGIYETITIDSETLQMSA 899
QY  482 IKKVI-----EKGYREKGOAIEYSGLTETQLAATOLAIIYFTD---SAIL 524

```

```

DB  900 SKKVTYTAPIITLKESGDEAGIYVYLAFTFYTHNVEDENQALAKVSEFELIDNVTHTATER 959
QY  525 DKDKLADYHGFQM-RNSTLAIAKILVEVAQDSNPQOLT-----DLDFRIP----- 569
DB  960 TTDEKQGYSDALMTGDYTLRYTNVPEYSVDEE--YLGGKAIRKLVGDQOLKIPLTKTI 1017
QY  570 NNNKYO-----SLGTQWHEPDLVDILIRMEDK-EVILPYTHNLTKRYTGLAGDRTKDF 623
DB  1018 DHSRLQYKDSITIVGDSMKPEE--NFVSATDKTGQDVPE-----KITVSGYDNKKAGV 1070
QY  624 HFEIELKNNKQELLISQVTKTKTNLEFKDGKATINLKHGSELTGLQLPBGYSYLVKEDS 683
DB  1071 YPIYSDEGKEETFAVYTVKPDQSKLEVKD--TTIYV--GDSWK-----PE--DNFVSATDK 1120
QY  684 EGYKV---KNSQEVANATYSKGIISDELAEEN 715
DB  1121 TGQDVPEPKIDVQGYVN--VDKITG---DYEIVKVN 1150

RESULT 7
ABP43430
ID  ABP43430 standard; Protein; 1416 AA.
XX
XX  ABP43430;
AC  XX
XX  05-AUG-2002 (first entry)
DT  XX
XX  E faecalis EF104 antigenic fragment.
DE  XX
XX  Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.
XX  Enterococcus faecalis.
OS  XX
XX  US2002045737-A1.
PN  XX
XX  18-APR-2002.
PD  XX
XX  04-MAY-1998; 98US-0071035.
PF  XX
XX  04-MAY-1998; 98US-0071035.
PR  XX
XX  (HUMA-) HUMAN GENOME SCI INC.
PA  XX
XX  Choi GH, Bailey C, Hromockyj A, Kunsch CA;
DB  WPI; 2002-425450/45.
XX  N-PSDB; ABN98186.
XX
XX  New genes and polypeptides from Enterococcus faecalis, useful as
XX  PT  vaccines for preventing, treating or attenuating an infection caused by
XX  PT  a member of the Enterococcus genus in an animal, particularly E.
XX  faecalis
XX
PS  Claim 9; Page 185; 255pp; English.
XX
XX  The present invention provides the protein and coding sequences of a
XX  CC  number of polypeptides from Enterococcus faecalis. The proteins can be
XX  CC  used as vaccines for preventing or attenuating an infection caused by a
XX  CC  member of the Enterococcus genus in an animal, particularly E. faecalis.
XX  CC  The polynucleotide is also useful for preventing or treating E. faecalis
XX  CC  infection. The present sequence is a protein of the invention.
XX
SO  Sequence 1416 AA;

Query Match 4.1%; Score 161; DB 23; Length 1416;
Best Local Similarity 21.1%; Pred. No. 0.017;
Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;

QY  39 AUYTSMVGAKTIVGLVES---STPNAINPDSSEYMYGYESVVRGHPPYKKORVAHDLR 95
DB  484 SUSTPYIGPKAIQVSDOTIEPIISVNP-LNLETAMGNDO-----NGAYSSR 531

```

QY	96	-VLEGSRSYQVYCFNLKKAFF-----LGSOSVKKWKKKHNDGISTREKEDYAMSRLTGD	149
Db	532	TVVSWGSKKEKPIQNLKKAHPYLSIRAKKEIYFYK-----LGT--DYVTPPTS	584
QY	150	ELNOKLAVVMNGHPONANGIMELE-----PLNAIRYTOAVVMYSDNAPISNPDESFK	204
Db	585	VIKRTPTITNEIQIPIGENVYPPDLPKDKSLPVDTIPITMAEGLTPDVTY----NSK	641
QY	205	RESESNLVSTSQSLMRQALKQLIDPLNATKMPKQVPDDFQLSIFESEDKGDYKNGYON	264
Db	642	RGSERTLQSSKNQGLVNARNDSPDLSVRFKIPAGA--DYLFEDYVSN--DQYDSIYQ	697
QY	265	LLSGGLVPTKPTPTGDDPMP--PNOPTSVLT-----KKYALGDYSKL--LEGATLQ	313
Db	698	YMDRGQYEDKRMPTNSPGYPTTPEDENTNSTYTFPGKTKKRYII-EYKNANGMWDVPTLY	756
QY	314	LTC-----DNVNSFOARVFNSSNDIGRIEISLDDTYTLTELSNPAQSVIAEPIFEKVAG	367
Db	757	ITGAKAPEQSNNGNSASVSYON---EALDLSLT-----QANPPLKANVTKT	801
QY	368	KVYT-IIDGK--QIENPKKEIYEPYVEAYVNDPEEFSVLTQNTAKAFYKANKNGSSQV	424
Db	802	TVTFTKNIDNKHRYKKNPTIELPRKGTNAQIDLNSIV-----KGPEDA	846
QY	425	YCFMADLAKSPDSDSGKTPRDTTGE---VKTHLAGRLFKYTKKPRDTPDTFLKH	481
Db	847	Y-----SLEKTYNGAKVIFKEDDTLENTIIEYNTVSANAGOJYETTTIDSELNOMSA	899
QY	482	IKKYV-----EKGYREKGOAIEYSGLTETQLRAQTOLAIYFTD-----SAEL	524
Db	900	SKKRYVTAPTTLTKSEGDAGEIYVLAATFTTHVEDEDNQIAKVSFELIDVTHLTATEF	959
QY	525	DKDKLMDYHGFQDM-NDSTLAVAKILVEYAODSNPOLT-----DLDFETP-----	569
Db	960	TTDEKGGYSDAIMTGDYTLRTNVPRQESYDDE--YLTGAIKILYKGNQDNOKIPDTKI	1017
QY	570	NNNKYQ-----SLIGTQWHPEDLDVDIRMEDKK-EVIPVYTHNLTKRTVTGLAGDRYKF	623
Db	1018	DHSHTAQVKDSTIYVGDGSMKPEE--NFVSATPDKTGQDVEF-----KITVSGQVNNXKAGV	1070
QY	624	HFEELKNNNOELLQVVKTKDKTLEFRKDGKATINLKHGESLTIOGLPEGSYLKVEDS	683
Db	1071	YPIIYDSEGEETRAYVTKPRQSKLEVKD--TTIYV--GDSWK---PE--DNEVSATDK	1120
QY	684	EGYKV---KVNQSEVANATVSKTGTISDETLAFEN	715
Db	1121	TGQDVPREKIDVQGTNV--VDKIG--DYELIVKYN	1150
RESULT 8			
ABU13709			
ABU13709	ID	ABU13709 standard; Protein; 1416 AA.	
XX	XX		
AC	ABU13709;		
XX	XX		
DT	26-FEB-2003 (first entry)		
XX	XX		
DE	Enterococcus faecalis EF040 polypeptide #202.		
XX	XX		
OS	Enterococcus faecalis.		
XX	XX		
PN	US6448043-B1.		
XX	XX		
PD	10-SEP-2002.		
XX	XX		
PF	04-MAY-1998; 98US-0071035.		
XX	XX		
PR	06-MAY-1997; 97US-044031P.		
PR	16-MAY-1997; 97US-046655P.		
PR	14-NOV-1997; 97US-066009P.		
PR	14-NOV-1997; 97US-066099P.		

XX	(HUMA-)	HUMAN GENOME SCI INC.
XX		
PA	Choi GH, Bailey C, Hromockyj A, Kunsch CA;	
PI		
PI	WPI: 2003-089120/08:	
DR	N-PsDB: ABX61756.	
DR		
XX		
XX	New EF040 polypeptides and polynucleotides from <i>Enterococcus faecalis</i> ,	
PT	useful for generating an immune response against <i>E. faecalis</i> and other	
PT	<i>Enterococcus</i> species, and as vaccines against other bacterial genera	
XX		
XX	Example 1: Column 207-208: 146pp: English.	
CC		
CC	The invention relates to polynucleotide fragments of a gene from	
CC	<i>Enterococcus faecalis</i> , EF040, and the polypeptides encoded by them. The	
CC	polypeptides are useful in detecting <i>E. faecalis</i> , as epitope tags, as	
CC	molecular weight markers on SDS-PAGE gels or for molecular sieve gel	
CC	filtration columns, in generating antibodies that specifically bind to	
CC	<i>E. faecalis</i> polypeptides, in generating an immune response against <i>E.</i>	
CC	<i>faecalis</i> and other <i>Enterococcus</i> species and as vaccines against other	
CC	bacterial genera. The polynucleotides are useful as probes for gene	
CC	mapping and for identifying <i>E. faecalis</i> in biological samples. Sequences	
CC	ABU13308-ABU13755 represent EF040 polypeptides of the invention.	
CC	Note: The sequence data for this patent can also be obtained from USPTO	
CC	at seqdata.uspto.gov/sequence.html .	
XX		
XX		
SQ	Sequence 1416 AA:	
	Query Match 4.1%, Score 161; DB 24; Length 1416;	
	Best Local Similarity 21.1%, Pred. No. 0.017;	
	Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;	
QY	39 ALVMSMGAKVVEGLVES--STFNAINPDSSSEYRWYGSYSYVGHPIYKQFVYADHLR 95	
DB	484 SLSTPVIQPKNAIQIOLVSDQYIEPISVNP-INAEIAMGNQDQ-----NGAYSSR 531	
QY	96 --VMESSRSRYOVYCFNLKAFP--LGSDSYKWKWKKKHGDISTKFEYDAMSRLTGD 149	
DB	532 TTVSVMSSKEKPIQNLKVKHNPVLSLRKREIYFYK--LGT--DYTVPTSDGS 584	
QY	150 ELNCKLRVMTNGHPQANNGIMEGLE---PLNAIRVTOEAWYVSDNAPISNDESEFK 204	
DB	585 VIKETPTITNEIQIIGENYVPDSLPKDKSIPVDIPIITMSAEGTLTVDVTYTT--NSK 641	
QY	205 RESSNLVSTQSLSLMQALQOLIDPNLATKMPKQVPPDPQSLFESDEKDKYKQVQN 264	
DB	642 RGSERRTQSSKNQELVNAANDSPSLVSKTIPAGA--DVLEDIYDVSN--DOVDSITPQ 697	
QY	265 LLISGLVTPKPTPGDDPMP--PNOPTVSVLI-----RKVAIGDYSKL--LEGATDQ 313	
DB	698 YMDGQYGFDRKMTNTNSPCYPTITTDENNTSTTPDEGKNNKXII-ETYNANGWIDVPLLY 756	
QY	314 LFG-----DNVNSFQARVSSNDIGERIELSDGTVLTLENSPAGYSIAEPTFFKVEAG 367	
DB	757 ITGTAKEPOSNNNGSASVQN--EALDILISAT-----QAAMPILKNVTKT 801	
QY	368 KVVY-IIDGK--QIENPKKELVEEYVSQAVYNDPEEFSVLTQONAKFYAKNNKSSQVY 424	
DB	802 TVTTKKNIDNKRKKNPTTELTLPKGTNAQIDLNSITV-----KGVPTDA 846	
QY	425 YCFNADKLSPDSEDSGGKTMTPDEFTTGC---VKYTHIAGRLDFKTVKPRDTPDEPTFKH 481	
DB	847 Y-----SLEKTTNGAKVIFKQDVTITNENITIEVNTVSANNGQIYETETTIDSETLNQMSA 899	
QY	482 IKKVI-----EKGYREKQAIEYSGLTETQRLAATQALAIYFTD---SAEL 524	
DB	900 SKKVTYAPRILTKFSEGDAGIEIVLATATFTYTHNVEDENQAIKVSFELIINVHTATFE 959	
QY	525 DKOKLKYHGGDM--NSTLAVAKIIVVEYADDSNPQLT-----DLDFITP----- 569	
DB	960 TTDEKGOYSPDAMTGTGDTLTKVTVNVPQETSYDEE--YLTGAKILVKQDNQOLKPLTKTI 1017	


```

XX  WPI: 2002-425450/45.
DR  N-PSDB; ABN98185.
XX
XX  New genes and polypeptides from Enterococcus faecalis, useful as
PT  vaccines for preventing, treating or attenuating an infection caused by
PT  a member of the Enterococcus genus in an animal, particularly E.
XX  faecalis
XX
XX  Claim 9; Page 182-183; 255pp; English.
XX
CC  The present invention provides the protein and coding sequences of a
CC  number of polypeptides from Enterococcus faecalis. The proteins can be
CC  used as vaccines for preventing or attenuating an infection caused by a
CC  member of the Enterococcus genus in an animal, particularly E. faecalis.
CC  The polynucleotide is also useful for preventing or treating E. faecalis
CC  infection. The present sequence is a protein of the invention.
XX
SQ  Sequence 1448 AA;

Query Match 4.1%; Score 161; DB 23; Length 1448;
Best Local Similarity 21.1%; Pred. No. 0.017;
Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;

OY 39 ALVTSMGAKTVFGLVES---STPNAINPDSSSEYRWGSESYVNGHYPYQGFVAHDLR 95
DB 511 SLSTPVIQPKAIQLVSDQYIEPISVNP-LNAETAMGNVDO-----NCAVSSR 558
OY 96 --VNLEGRSYOVYCFNLKKAFF---LGSDSYKWKYKKHKGISTKFEDYAMSFRITGD 149
DB 559 TTIVSMGSKKEPIQNLLEIKVHPNPLSLRAIKREIFYKK---LGT---DVTVPITSDGS 611
OY 150 ELNOKLRAVMTNGHPQANANGIMEGLE-----PLNAIRVTOEAVWYYSNAPISNDESFK 204
DB 612 VIKFTTPTINEIQIPIGFNVPDLSLKDKSIPVDIPIITMSAEGILTPTVDTVT---NSK 668
OY 205 RESENLTSTQSLSMRQALQOLIDPNLATKMPKQVDPDFQLSIPESDEKDGKVKYKQYON 264
DB 669 RGSERTLOSSKNQPLVNAKRNDSFDSLSVTKIPAGA--DVLEFDIYVSN--DQVDSITPQ 724
OY 265 LLSGLVLPKPTPTGDDPMP--PNOPQTSVLI-----RKYAIGDYSKL---LEGATLQ 313
DB 725 YMDGQYEDKPMTPRSPGYPITFEDENSTYTFDGGKTNKRITII-EYKNAWGMIDVPTLY 783
OY 314 LTG-----DNVNSFQARVFNSSNDIGERIEISDGYITLTLELNSPAGISIAEPIPKVYAG 367
DB 784 ITTGAKEPQSNNGSASVYQN---EALDILISAT-----QAAPTLKNTYKT 828
OY 368 KYVT-IIDGK--QIENPKIEIPEYSYVANDFEESVLTQNVAKFYAKKNKNGSSQGV 424
DB 829 TVTTKNIDKTHRKNPTELPRKGTNAQIDLSITY-----KGVPEDA 873
OY 425 YCFNADLKSPDSEDEGKTKMPTPTTGE--VKYTHIAGRLFKYTVKPRPTDPTFLKH 481
DB 874 Y-----SLEKTINGAKVIFKDYTLTENITIEYNTVSANAGQIYETETISETLNMOSA 926
OY 482 IKYVI-----EKGYREKGAIEYSGLTENQLRAATOAIYFPD---SREL 524
DB 927 SKKVVTAAPITLAKSEGDAGIEYVILATATFTTHNVEDNQIAKVSFELIDNVTHTATEP 986
OY 525 DKDKLDVHGFGDM-NDSTLAVAKILVEYAODSNPOLT-----DLDFEIP----- 569
DB 987 TTDERGOVSFPAIMTGDTLATVNVPOEYSDDE--YLTGKAIKLYKDNOLKIPLTITI 1044
OY 570 NNNKTYQ-----SLIGTOMHPEDLVADIIMEDKK-EVIRPTNHLTKRTVTGLAGDRTKDF 623
DB 1045 DHSRLQVSDSTIYVGDSSWKPEE--NFVSATDKTGQDVPFE-----KITVSGQVNDXKAGV 1097
OY 624 HFELEIKNNKQELLSQYFKTDKINLEFKDGKATINLKGESLTLQGLREGSIYLYKEDS 683
DB 1098 YPILYSDEGKEKETIYVYKPKPOSKLEAVD--TTIIV--GDSWK---PE--DNFVSATDK 1147
OY 684 EGYRV---KVSQEVANATVSKTGITSDTLAFEN 715

```

```

DB 1148 TGQDVPEFKIDVQGTIVN--VDKIG---DYELIYKN 1177

RESULT 11
ID ABU13708
ABU13708 standard; Protein; 1448 AA.
XX
XX  ABU13708;
XX
XX  26-FEB-2003 (first entry)
XX
DE  Enterococcus faecalis EF040 polypeptide #201.
XX
KM  EF040; immunostimulant; antibacterial; gene mapping.
XX
OS  Enterococcus faecalis.
XX
PN  US6448043-B1.
XX
PD  10-SEP-2002.
XX
PF  04-MAY-1998; 98US-0071035.
XX
PR  06-MAY-1997; 97US-044031P.
PR  16-MAY-1997; 97US-046655P.
PR  14-NOV-1997; 97US-066009P.
PR  14-NOV-1997; 97US-066009P.
XX
PA  (HUMA-) HUMAN GENOME SCI INC.
PI  Chol GH, Bailey C, Hromockyj A, Kunsch CA;
XX
DR  WPI: 2003-089120/08.
XX
DR  N-PSDB; ABX61755.
XX
PT  New EF040 polypeptides and polynucleotides from Enterococcus faecalis,
PT  useful for generating an immune response against E. faecalis and other
PT  Enterococcus species, and as vaccines against other bacterial genera
XX
XX  Example 1; Column 205-206; 146pp; English.
XX
PS  The invention relates to polynucleotide fragments of a gene from
XX  Enterococcus faecalis, EF040, and the polypeptides encoded by them. The
XX  polypeptides are useful in detecting E. faecalis, as epitope tags, as
XX  molecular weight markers in SDS-PAGE gels or for molecular sieve gel
XX  filtration columns, in generating antibodies that specifically bind to
XX  the E. faecalis polypeptides, in generating an immune response against E.
XX  faecalis and other Enterococcus species and as vaccines against other
XX  bacterial genera. The polynucleotides are useful as probes for gene
XX  mapping and for identifying E. faecalis in biological samples. Sequences
XX  CC ABU13508-ABU13735 represent EF040 polypeptides of the invention.
XX  CC Note: The sequence data for this patent can also be obtained from USPTO
XX  at segdata.uspto.gov/sequence.html.
XX
SQ  Sequence 1448 AA;

Query Match 4.1%; Score 161; DB 24; Length 1448;
Best Local Similarity 21.1%; Pred. No. 0.017;
Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;

OY 39 ALVTSMGAKTVFGLVES---STPNAINPDSSSEYRWGSESYVNGHYPYQGFVAHDLR 95
DB 511 SLSTPVIQPKAIQLVSDQYIEPISVNP-LNAETAMGNVDO-----NCAVSSR 558
OY 96 --VNLEGRSYOVYCFNLKKAFF---LGSDSYKWKYKKHKGISTKFEDYAMSFRITGD 149
DB 559 TTIVSMGSKKEPIQNLLEIKVHPNPLSLRAIKREIFYKK---LGT---DVTVPITSDGS 611
OY 150 ELNOKLRAVMTNGHPQANANGIMEGLE-----PLNAIRVTOEAVWYYSNAPISNDESFK 204
DB 612 VIKFTTPTINEIQIPIGFNVPDLSLKDKSIPVDIPIITMSAEGILTPTVDTVT---NSK 668

```


Db 1297 -LTKLDSKSRSLAGAEFEELQTKLGSLKDKVVT-EANGQLQIDNLAPGQYLVETKAPT 1354
 QY 551 EYADQSNPQLDLDFEIPNNKKYQ-----SLIGQWHP-----DLV 588
 Db 1335 GVDLADTP-----VEETLEPNOKAPIQVTKTNMSTGSAVLTKGTGKALLANATFKLV 1409
 QY 589 DIIRMEDKKEVIVPYNHLNL-----RKVTGLA-GD-----RTK-----DFHEIELKN 631
 Db 1410 D-----EDNN-----VTENLTTDASGKLEITNLAPGYQLETKAPAGYELDTVPVVKITF 1461
 QY 632 NKQELLS-----QTVKTDKTNLEFKDKATINLKHGSLVLQGLPEGYSLVYKETTDEG 685
 Db 1462 DQKETLQVTKTNLKTIVSGKVAEFDVTKGNV-LAEKEIHT--GI-VGDKYATKAKADIKG 1517
 QY 686 YKV-----KVSQEVANATVSKTGITSDFLAPENKPEPV----- 721
 Db 1518 YKLTQPTNKGSEVFETEDQK-----TFVEKNKAPIVVNDKPVTPVKKTPVPAKK 1571
 QY 722 -----PTGVDQKINGYALALIVAGISLGIVGHTIRIK 755
 Db 1572 PTVKTSLPSTGDESPYG-----IIFGLFASFMGLFLKRSK 1608

 RESULT 13
 ID ABB47329 standard; Protein: 1530 AA.
 AC ABB47329;
 DT 05-FEB-2002 (first entry)
 DE Listeria monocytogenes protein #33.
 XX Listeria monocytogenes protein #33.
 KY Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KM vitamin B12; bacterial infection; disease.
 OS Listeria monocytogenes.
 PN W0200177335-AZ.
 PD 18-OCT-2001.
 XX 11-APR-2001; 2001WO-FR01118.
 PF 11-APR-2000; 2000FR-0004629.
 PR (INSP) INST PASTEUR.
 PA
 XX Buchrieser C, Frangeul L, Couve E, Rusnlok C, Fsihi H, Dehoux P;
 PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Dommann E, Hahn T, Berche P, Chabdit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Madeno E, De Pablos B, Wehlund J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX
 DR WPI: 2002-010914/01.
 XX
 PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides
 XX
 PS Claim 6; SEQ ID No 34; 192pp; French.
 CC
 CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and

CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1530 AA;

Query Match 4.0%; Score 156; DB 23; Length 1530;

Best Local Similarity 20.6%; Pred. No. 0.045; Indels 316; Gaps 38;

Matches 160; Conservative 78; Mismatches 221; Indels 316; Gaps 38;

QY 193 NAPISNPDESEKRR--ESESNLVSTQSLIMROALKQIDPNLTKAPRQVPPDQSLIFE 250
 Db 794 NMPISPGYEDLKTGYDESN-----LEFYQKFNELNQSIVIK-----YQTAITL 839
 QY 251 SEDK-----GDKYKGYQNLISGLVPTK-----PPPGDPPMPNPQPTTSVLIRKAI 300
 Db 840 TSDTETTAQIGNSVFTGDNITKGETETKNIETKIGD--GTGETGKIILNKYDK 896
 QY 301 GDYSKLLGATLQLGDNVNSFOARFSSNDIGETIELSDGYTLTLELNSPAGYSI---- 356
 Db 897 ADPSILPGATFDLIA--DDEKVDYQTTDKNGVIEDDLVYGDYTLKEVASAPGATPLPAS 955
 QY 357 AEPITKYE-----AGKVTYIID--GKQIEN 380
 Db 956 TENIVQKLEODEKVVQVWNNKMKPIKETGEVHLVTKDTKATGATLAESISLXKSAEGLN 1015
 QY 361 -----PKKEIV-----EPYSVAIVDEE--EFSVLTQNTAKFTYAKN--- 417
 Db 1016 GLTTDENGELTIHNLDSYVLYKTRKAPGKYLSEKTEFVSQGVDAIEIOAENEDL 1075
 QY 418 -----NGSSQV-----YCF----- 427
 Db 1076 GEAVLTAKDSETNALSGAKFNLNDSGEVIOITNLVSDENGEIRVQNLPEGDYAFQETEA 1135
 QY 428 --NADLKS-----PPDSBDG-----GKT----- 443
 Db 1136 PTNYDLATNTMFTLIVAGQTSATWVTAENNKTKGPDVDTGEVILVKQDSAGGETLEGAVF 1195
 QY 444 --MTPD-----FTT-----GEVYKTHIA-----GROLFTYVAPROTD 474
 Db 1196 DLTMDGAIIVASNLTTDANGELITVNLAPGKYSFKETRAPEGYELATVWETIAPNPOE 1255
 QY 475 -----PDT--FLKHKKVIRKGYREKQALEYSGLTQTOLRAATQALYYFT 519
 Db 1256 KITTAENMTKLAIPIDPDAASVKITKQDSENGVRLAG--AEFSLIAENGELQTNLK----T 1309
 QY 520 DSAELDKDLKLDYHGFQMDNSTLAVAKILVEYADQSNPQITDLDFEIPNNKKYQSLIG 579
 Db 1310 DEA-----GELEVNNLAPGVNRI--QETKAP-----DGYQ--LES 1340
 QY 580 TQMHPEDLVDIIRMEDKKEVIVPYNHLNLKRYV--TGLADRTKPFHEIELKNNKQEL 637
 Db 1341 TPWQPE-----IVANDTSQVYVIAENALPEPVATGAVRLIKTDTSETGTRLSGAVSFL 1395
 QY 638 SOTVTKDTNLEFKGKATINLKHGSLVLQGLPEGYSLVYKETS--BGYKV----- 688
 Db 1396 DESGVQIANLTDE-----KE--IFIDCLTIG--NYSIAKETVAPPGCYELAEQPMWQ 1445
 QY 689 -----KVSQEVANATVSKTGITSD-----ETLAFENKPEP 719
 Db 1446 IVKGVDAVVIKAENSPITANGAISFEQEDTDKPSIEIPVKTDTLATETVTKLP 1500

 RESULT 14
 ID ABB47336 standard; Protein: 832 AA.

XX ABB47336;
 AC 05-FEB-2002 (first entry)
 DE Listeria monocytogenes protein #40.
 XX
 XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KM vitamin B12; bacterial infection; disease.
 XX
 OS Listeria monocytogenes.
 XX
 PN W0200177335-A2.
 PD 18-OCT-2001.
 XX
 XX 11-APR-2001; 2001MO-FR01118.
 PF 11-APR-2000; 2000FR-0004629.
 PR (INSP) INST PASTEUR.
 XX
 PA Buchrieser C, Frangoul L, Couve E, Rusniok C, Eshti H, Dehoux P,
 PI Dussurget O, Chtouani F, Nedjari H, Glaser P, Kunst F, Cossart P,
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA,
 PI Dominguez-Bernal G, Garrido-Garcia P, Tlerriz-Martinez A, Amend A,
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L,
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
 PI Muedeno E, De Pablos B, Wehlend J, Kaerst U, Entian K, Hauf J,
 PI Rose M, Voss H.
 XX
 DR WPI; 2002-010914/01.
 XX
 PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides
 PS Claim 6; SEQ ID No 41; 192pp; French.
 XX
 CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes BGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by L.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 832 AA;
 SO
 Query Match 3.9%; Score 155; DB 23; Length 832;
 Best Local Similarity 19.9%; Pred. No. 0.02;
 Matches 141; Conservative 93; Mismatches 232; Indels 244; Gaps 33;
 QY 155 LRAVMNGHPQNNAN-GIMEGLEPLNARVQEAVMYSDNAPISNPDESFKRESESTLV-212
 DB 176 LKSIANNNNKTTGNFSIVKLPPLHTLEV-----LGNATTEIDINQPLVTV-222
 QY 213 -STSOISIMROALKOLIDPN-----LATKMPKOVDPDDQLSJFESE-252
 DB 223 LSNDELLEKTLTKLNLSQNGLGRIASSISIDMGDLESYVLMLPEIISYDISGNVLDSD-282
 QY 253 DKGDKNKGYONLLSGLVPTKPPPTGDDPPMPNPQOTSVLIRKYYAIG--DYSKLL--307

DB 283 DIHLEMLPAVKNNLDISSENLRLPKINDPEPL-----LTTINVRSNKKIDRLSSKLVDPV-336
 QY 308 -----EGATLQLT-----GDNVNSFQA-----RVFSNDI-----GER-335
 DB 337 KLATLNADKQAVTLSTKIIAAGNFTIPNNVENLAGQWVTPKIIISNGYSQSIAMASGEL-396
 QY 336 IELSDGTYLTLE-LNSPA-----GYSIAEPTFEKVEAGKYVYTIIDKQLE-----379
 DB 397 SGLSKVSYTFDEYVINSALIGKXTGYVNPQIEVKA-----VPYVAKSVSYAPVANKDEA-452
 QY 380 -----NPKIEVPEYSVEAYNDEEESVLTQNAKFEYVANKNGSSOVVYCFNA-429
 DB 453 TFLQDIRASASENAQITSQDS-----EVYDEATFGDYVTVLHAKNE-----F-494
 QY 430 DLKSPDSEGGCTMPDFTTGEVKYTHIAGRLDFKYVPRPTDDPTFLKHKKYIENG-489
 DB 495 DLKA-----DPYTVVYHINDIQKPO-514
 QY 490 YREKGAIEY---SGLTETOLRA-----ATOL---AIYPTDSAEIDKDKLKDYG---534
 DB 515 VAVNSNDISFEVGETELTSEVLAKSGAVYTDLDEAIKMEVDLSKLTGYEFTIA-574
 QY 535 ---FGDMNSTLAVAKILVEYADSNPP---QLTDLDFIPNNKKYQ-----SLIGQ-581
 DB 575 KSKSGASDPKILSVKIV-----DTEKPIIQINNPEIIKSGELTEGOIIDQGITARD-629
 QY 582 WHEPEDI---VDIIRMEKKEVIVTNTLTKRTVYGLAQRTPDHPFETELKNNKOLL-638
 DB 630 NYDODLNIHMDLSKVDISK---PGSYEVTV-----YTEDSSGKSELYV-670
 QY 639 QTVKTDKTNLEFKGKATI-----NLKHGESLTLQD-LPEGYSVLKETSSEGYKVVN-691
 DB 671 ITVAKVEPARI---GRITTIYMOSENNELAESNTITGEVGETYETLAKEL--EGYTLKEN-724
 QY 692 SQ-----EVANATVS---KTGITSDETLAEFNKKEPVYPTGVQDKING-731
 DB 725 PANSGVFEETROTIOYIVKDIINPEFVYSNNVTPELPSNNNSVNG-774
 RESULT 15
 AAM17900
 ID AAM17900 standard; Protein: 1849 AA.
 XX
 AC AAM17900;
 XX
 XX 29-JAN-1998 (first entry)
 DE Photorhabdus luminescens insect toxin TcdA11.
 XX
 XX Insecticide; insect; pest control; biological control;
 KW Photorhabdus luminescens; TcdA; Southern corn rootworm;
 KW Colorado potato beetle; Western corn rootworm; meal worm;
 KW boll weevil; turf grub; Coleoptera; beet armyworm; black cutworm;
 KW cabbage looper; codling moth; corn earworm; European corn borer;
 KW tobacco hornworm; tobacco budworm; Lepidoptera; Hymenoptera;
 KW Diptera; Dictyoptera; Acarina; Homoptera.
 XX
 OS Photorhabdus luminescens strain W-14 (ATCC 55397).
 XX
 FH Key Location/Qualifiers
 FT Protein 1..1849
 FT /label= TcdA11
 FT Peptide 1..12
 FT /note= "S2 N-terminus (Claim 30)"
 FT Peptide 196..211
 FT /note= "Cryptic peptide (Claim 30)"
 FT Peptide 466..475
 FT /note= "Cryptic peptide (Claim 30)"
 FT Peptide 993..1004
 FT /note= "Isolated N-terminal peptide (Claim 30)"
 FT Peptide 1297..1312
 FT /note= "Cryptic peptide (Claim 30)"

